BEST PRACTICES FOR WATER QUALITY TRADING
JOINT REGIONAL AGREEMENT

Discussion Guide, June 5-6th, 2013

This Discussion Guide is intended to provide definitions, context, analysis, and options for addressing various components of water quality trading programs (e.g. trading ratios, BMP quality standards). It poses questions that will be discussed at the interagency workshops. This document may reference other trading programs, examples, or documents, but is not intended to serve as a published report or white paper and thus will not be extensively cited. This document will be included in the workshop packet and posted online following each workshop.

5. Credit Quantification at individual project sites; 3. Project site assessment

Credits for pollution reduction are calculated by subtracting the pollution load associated with the anticipated future conditions from the load associated with current conditions. Quantifying credits relies on the best available science to predict and/or measure the pollution reduction from BMPs implemented. Generally, credit quantification looks at three scales:

- Estimates of pollution reduction <u>at the edge of field</u> from BMP implementation;
- <u>Delivery</u> of that pollution reduction from the edge of field into the nearest stream; and
- <u>Attenuation</u> of that reduction from the nearest stream to a downstream point of compliance.

Each of these scales involves different dynamics between soils, hydrology, and land management, often necessitating the use of different tools or models. This memo describes the characteristics needed for models/quantification tools to meet the needs of trading at each scale. The memo also discusses a process for selecting, calibrating, validating, and approving particular models for specific trading programs. Section numbers correspond to the organization of topics in the Tier 2 draft outline.

5.0 a Options for quantifying water quality benefits for trading

Quantification methods can be grouped into three general types: modeling, pre-determined rates and ratios, and direct monitoring. The following was adapted from *In It Together* (Vol 2, p.20).

Modeling: Many programs use modeling in their approach to water quality trading. Where existing models can suit program needs, and where sufficient local data is available for calibration and validation, models can be more site-specific than using standard rates.

Predetermined rates: This approach involves setting standard values for water quality benefits based on the best available science. These values are often expressed as ratios/percentages (e.g. 50% of the phosphorus load will attenuate between points A and B) or absolute loads (e.g. use of cover crop will reduce sediment loading by 35%). Some rates are grounded in extensive research and modeling, while others are adapted from relevant literature. BMP effectiveness rates provide a high level of repeatability and predictability in a trading program, but they are not as sensitive to site-specific conditions as modeling approaches. Start-up costs to generate these rates may be high where relevant studies or modeled values are not available, but the cost of maintaining the approach over time is likely to be low. The costs would be more associated with obtaining the long-term data necessary to evaluate attenuation rates or absolute load reduction.

Direct monitoring: Direct measurement is often used for ambient water quality monitoring at the reach or watershed scale and serves as an important tool for calibrating and validating models but is not typically used in trading programs for quantifying water quality credits, largely because it is the most costly measurement system to implement. Direct monitoring may be appropriate in those cases where there are a limited and/or controllable number of variables affecting water quality (e.g. improvements across an irrigation district where inputs and outputs can be closely monitored) and where clear monitoring and quality assurance plans are in place.

5.0b Identifying, evaluating, and approving quantification methods

Several regions/states have developed processes for review and selection of scientific methods in general. The process will differ depending on which type of approach is being used for a particular water quality credit in a given phyisographic location.

Modeling: Selection and review of modeling approaches may occur by 1) identifying methods that fit the intended uses, users, and evaluation criteria, 2) adaptation to local conditions, 3) technical review, and 4) formal approval.

1. *Identify relevant methods*: At the most basic level, a model will need to deliver outputs in useful units. For water quality trading, this means expressing model outputs in the same units as the regulatory water quality standard, or their surrogate targets, typically load (e.g. lbs) on a timescale that is monthly or finer scale (e.g. allowing for seasonal outputs that can correspond with seasonal load limits). A model will also need to operate at an appropriate scale and resolution. Edge-of-field and delivery-from-field-to-water models will typically need to work for a 1–3 acre field or a 300–3,000 acre field. Attenuation models should operate at

the reach- to watershed scale and be able to accommodate multiple inputs and outputs to better reflect cumulative patterns and loading processes.

Beyond these basic requirements, a good model for water quality trading should be:

- Accurate: representing true pollution load reductions. Assessments of uncertainty around the estimates provided can help to represent the level of accuracy;
- Repeatable: if two people apply the model with the same data in the same place, they get the same answer. Protocols or user guidance can greatly improve the consistency of model application;
- Sensitive: variation in model output reflects actual variability in the water quality indicators being measured, and not random noise produced by the model;
- Transparent: easy to understand indicators and well-documented relationship of inputs and indicators to an overall estimate of pollution reduction. The model should be well vetted in the scientific community and also be posted in the public domain for use by anyone without charge.

A good field-scale model will also be practical and economical to set up and apply, easy to use for the targeted user group, and compatible with other relevant models so that its outputs can plug easily into reach or watershed scale models and evaluations of overall program performance.

Depending on available resources, one or more tools that meet these criteria are selected for calibration and validation. Finding the perfect model that meets all of these criteria may be difficult in many situations. Almost always, trading programs will need to make some tradeoffs in which models they select and orient their selection based on the program's objectives. For example, models that are more complex may more accurately represent the dynamics driving water quality changes, but that complexity may also make them harder to use and less transparent.

2. *Adapt to local conditions (Calibrate)*: Calibration includes changing model parameters to better match local conditions. Ideally, calibration occurs using measured water quality data from various locations in the watershed, including project sites. Part of calibration may also mean developing and integrating standard datasets for the local area (e.g. soils, climate, and crop management) or altering the coefficents of certain model parameters based on expert judgement.

3. *Technical review (Validate):* Validation tests whether the model outputs or other methods are meeting evaluation criteria (accurate, repeatible, sensitive, transparent). Often, validation include comparison of model results with measured data, sensitivity analyses, and uncertainty analyses. Validation may also include comparision with other model outputs, literature values, or expert judgement. Where measured data is not available to validate accuracy, adapative management and monitoring to improve the model over the time is particularly important—see the section below on adapative management. An analysis to describe model uncertainty is also part of validating accuracy, and can enter into the process of determining trading ratios.

Model validation may be an internal process or conducted by an independent entity. In either case, results of the technical review should be made publicly available and incorporated into technical documentation as possible. This could include publishing of results in peer-reviewed scientific literature.

4. *Formal Approval*: Formally approving the model or tool might come in the form of inclusion of the tool within state guidance, an approval letter from the state water quality agency or EPA, or approval to use the tool within an NPDES permit.

Pre-determined rates: To use literature values or other means as the basis of standard rates, justification for those rates should include documentation of how the values were arrived at, why those rates are appropriate for/transferable to the proposed trading geography and conditions, and some guidance and analysis about the likely sources of variation in performance of those BMPs based on local conditions. State agencies would perform a technical review and formal approval similar to that described for modeling approaches.

Direct monitoring: Direct monitoring may be appropriate in those cases where the project developer can have "control" over enough of the factors shaping water quality to show a measurable improvement in water quality. Some of the most likely BMPs/situations where direct monitoring may work include improvements across an irrigation district where inputs and outputs can be closely monitored in one or a set number of ditches and drains. To use direct monitoring for quantifying credits, there needs to be a clear monitoring plan and quality assurance plan approved by the state agency or their designee. The project developer needs to use instrumentation that captures water quality samples at enough frequency to create an estimate of average water quality improvement over a specified time (e.g. year, season, or month) and produce estimates of variation within that time period. That instrumentation needs to be objectively verifiable—a verifier can confirm that the instrument is appropriate for the purpose, installed and calibrated correctly, and producing adequate results. Records need to be kept for each sample taken, including date, time, method of data collection, and results. State agencies would perform a technical review and formal approval of the monitoring plan and data acquired similar to that described for modeling approaches.

5.0c Adaptive management of quantification methods

Quantification tools can always be improved, and sometimes the best way to improve them is to use them. In addition to monitoring that projects are in place and meeting quality and performance standards, trading programs should make sure that a representative subset of projects are gathering the data needed to improve quantification tools over time. This might include asking a sample of projects to install direct measurement devices to gather data needed to validate the quantification tools. For nutrients, needed data might include various types of water and soil samples, and flow discharges. For temperature, this might include characterizing shade-generating features on the project site (e.g. riparian vegeation type), measuring effective shade and taking upstream and downstream temperature measurements. This data would not be used to determine compliance for the project developer or NPDES permittee, but would be used to improve the program.

II. Recommended default: Trading programs should use the best available methods to quantify credits. In many cases modeling approaches used in TMDL development should be used to support trading where possible. Where the science and measurement capacity is already available to support pre-determined pollution reduction rates or direct measurement of reductions, those methods should be used. Where these approaches are not sufficient to quantify pollution reductions from individual projects or attenuate those reductions from points of generation to points of compliance or concern, trading programs should have standard methods for quantifying credits that have been formally approved by the relevant agency. Quantification methods should be repeatable, sensitive, accurate, practical, and transparent. Methods that have a longer history of usage and application and a documented track record would be preferred. Documentation of approved methods should include a thorough review of technical basis, procedures for consistent application, and a plan for improving the method over time. Methods and associated documentation should be publicly available, and, where feasible, vetted through a public- and peer-reviewed process.

III. Reasons to deviate from the default: Where standard quantification methods are inappropriate or insufficient, project developers may also have the opportunity to develop project-specific water quality benefit or load reduction estimates. Review of these projects will require significant effort by agency staff, and is most appropriate for large projects that will require substantial design and review, and will generate substantial water quality benefits. Project-specific methods must meet the same standards applied to program-approved models and tools.

5.1/5.2.a Quantifying conditions at the edge of the field

Modeling: Many programs are moving toward modeling water quality improvements to quantify credits. There are a couple ways to do this.

> *Site-Specific Calculations* consider farm-specific variables like soil type, historic rainfall, slope, prior cropping patterns, and crop management data to produce estimates of baseline and post-action nitrogen, phosphorus, sediment, crop yield, and flow at the edge of a field. These models can be complex because they require data for all those variables, as well as methods for calibration and validation. The Agricultural Policy/Environmental eXtender (APEX) is one model capable of performing site-specific calculations, with routines for calibration and validation. Tools like NutrientNet and USDA's Nutrient Tracking Tool (both user-interfaces for APEX) seek to reduce the complexity of performing these types of calculation by automating the process of identifying and integrating site-specific soil and climate data.
>
> *Custom Calculations* use average data for all agricultural operations in the land area under a program (e.g. using an average rainfall or average soil type for a watershed). This simplifies calculations and reduces program administration because regulators do not need to make determinations on modeling and is sufficient for environmental protection as long as the averages are conservative. However, the accuracy and sensitivity of the calculations may suffer.

Either type of model will require some level of local calibration or adjustment because factors like erosion rates, hydrology, and farming practices differ across parts of the United States. Models can be more sensitive to site-specific conditions than standard BMP effectiveness rates, but not necessarily more accurate when those BMP effectiveness rates have been well researched (e.g. in the Chesapeake program). Calibration of models requires reliable data on BMP effectiveness to generate accurate water quality estimates. If the data or resources are not available to apply an existing model, program designers may need to start off with a different approach (e.g. using pre-determined BMP efficiencies discussed below).

Pre-determined rates (BMP effectiveness rates): Most of the early water quality trading programs used BMP effectiveness rates to quantify water quality improvements. Ideally, BMP effectiveness rates are not used in isolation, but are used as inputs and companions to other modeling approaches to quantify credits. Furthermore, BMP effectiveness rates ought to be evaluated over time to ensure they are meeting their stated objectives and water quality targets.

Direct monitoring: Hydrologic systems are highly dynamic, so directly measuring in-stream changes from one BMP, in one field, in one season is very difficult. It can also be expensive, requiring the use of water quality monitoring equipment, ground or airborne fieldwork, deliberate study design, and statistical analysis. Direct measurement may be more viable for some BMPs or agricultural systems, for example, where discharge flows through ditches or canals and is easily measurable or where proxies can be determined through remote sensing (e.g. canopy closure and density of a riparian reach)

I. Options and examples

<u>What are some of the field scale tools in use and/or available for use in trading in the region?</u>

> *Nutrients:* Hydrologic characterization tool (developed by University of Idaho); Nutrient Tracking Tool; BMP efficiency rates (e.g. the ones explored for Spokane); STEP-L; APEX
>
> *Sediments*: Surface Irrigation Soil Loss (SISL) model; Hydrologic characterization tool (developed by University of Idaho); STEP-L ; streambank erosion inventory (Idaho); Watershed Erosion Predition Project (WEPP); Revised Universal Soil Loss Equation (RUSLE).
>
> *Temperature*: HeatSource modules and extensions--Shade-a-lator and the T-Tools extension (OR, ID); Shade (WA, similar to shade-a-lator); CE-QUAL-W2; HEC-RAS; Potential Natural Vegetation (PNV) shade analysis; W3T to quantify temperature benefits of in-stream flow (in development by National Fish and Wildlife Foundation)

What is missing? What are other models or methods for quantifying pollution reductions at the field scale that you would like to learn more about?

II. Recommended default: Several options exist, does it make to set a regional default for field-scale quantification methods or is the process/criteria described above sufficient to cover those situations where the model from the TMDL cannot be used?

III. Reasons to deviate from the default: Unclear. Are there specific instances you can imagine where a trade would require different technical support than both the TMDL models and methods approved through the process above?

5.1/5.2.b Delivering pollutants from the edge of field <u>into the stream</u>

Not all farm fields flow directly into a stream, and not all pollutants will transfer from the edge of a field into the nearest water body. In some instances, trading programs have assumed that 100% of pollutants leaving the edge of a farm field adjacent to stream reach the water column. Other trading programs have used delivery ratios to determine the percentage of pollutant that reaches a stream. A growing number of programs are now using attenuation models to quantify the delivery of pollutants into a stream and between points in the watershed.

I. Options and examples

<u>What are some of the delivery tools in use and/or available for use in trading in the region?</u>

There are not a lot of models that incorporate delivery from the edge of a field into a stream. When they do, the models are usually operated at the reach or watershed scale. Missing these models, some trading programs have A) used a standard delivery factor (e.g. Boise program), or B) required that all credits come from fields immediately adjacent to a stream (e.g. Willamette Partnership) as a surrogate assumption for delivery of all nutrients into a stream.

Several trading programs have developed delivery factors, the methodology for which may be transferable to the Northwest.

II. Recommended default: A calibrated and validated model or an approved delivery factor based in science is preferable, but a transparent surrogate for delivery such as location alongside a stream or other permanent water body may be considered.

III. Reasons to deviate from the default: Delivery modeling or ratios may not be necessary for BMPs relating to irrigation systems where the hydrologic connection between the discharge water and receiving water body is direct or nearly so.

5.1/5.2.c Delivering pollutants <u>through the watershed</u>

Most often, reach- to watershed-scale models will be used to attenuate a pollutant through a watershed down to a point of compliance. Reach scale quantification is almost always based on models, often the same models that were used to develop the TMDL in a watershed. In some cases, either where there is no TMDL yet or where a TMDL is not sensitive enough to attenuate load reductions from a smaller nonpoint source, other reach scale models need to be used.

I. Options and examples

<u>What are some of the reach and watershed scale tools in use and/or available for use in trading in the region?</u>

Nutrients: QUAL2K, QUAL2Kw, CE-QUAL-W2 and flow duration curves have been used in many nutrient TMDLs. Their ability to attenuate nutrients for trades is unclear. Other watershed models used or considered for use in trading include: WARMF, Better Assessment Science Integrating point & Non-point Sources (BASINS), Soil and Water Assessment Tool (SWAT), and the MIKE model suite can be considered for nutrient dynamics quantification..

Sediment: BASINS, SWAT, and the MIKE model suite also can be considered for sediment mobilitation and transport quantification.

Temperature: Heat Source, HEC-RAS, CE-QUAL-W2, W3T to quantify temperature benefits of in-stream flow for small reaches (in development by National Fish and Wildlife Foundation)

What is missing?

II. Recommended default: Several options exist, does it make sense to set a regional default for field-scale quantification methods or is the process/criteria described above sufficient to cover those situations where the model from the TMDL cannot be used?

III. Reasons to deviate from the default: Are there specific instances you can imagine where a trade would require different technical support than both the TMDL models and methods approved through the process above?

3. Project site assessment

This section discusses how to develop and document the information necessary to input into the models discussed in section 5, including the data and documentation necessary to establish current conditions on a credit project site and to forecast future conditions to predict water quality improvements.

3.1 Pre-project site conditions assessment

To quantify credits, a project developer needs access to information about on-farm operations and current conditions. There is also a need to compare current conditions with recent past activities to help establish baseline conditions. For example, annual practices like crop cover may change year over year, so there may be a need to look back over the last 3-5 years to establish a "baseline" pattern of crop rotation. A farmer may also need to document how much fertilizer has been used each year over the last three years or more to establish typical application rates and usage patterns. Each of these pieces of information is an input into models that will later quantify credits. A trading program also needs to clarify what documentation is required for each of those model inputs. Is it enough to have project developers attest that the information is accurate? Does there need to be specific monitoring for some types of information? What other kinds of records would be needed and/or reasonable?

I. Options and examples

<u>What documentation is needed to establish current conditions for specific BMPs?</u>

The information required to document current condition will vary depending on both the BMPs being proposed for credits and the type of pollutant credit being targeted. Below are some samples of information and documentation required for specific BMPs.

The Ohio River Basin program requires the following records for **three years** of farm practice history:

• Crop rotation sequence

• Crop residue management − Each crop within the rotation for each field; Yield/acre/ year and units, date of planting, date of harvest and whether residue is removed from field; If a perennial hay crop is grown, provide typical seeding date, number of cuttings and yield/acre; For tree crops, provide mo and year of establishment.

• Field operations− Provide tillage information for each field including equipment used, soil penetration depth, and type of residue managers.

• Crop nutrient input− Provide field identification, crop and yield goal, date of application, formulation of material applied, method of application, and actual lb/ac of actual N, P, and K that was applied.

• Irrigation water management (if BMP involves tile drainage) − Tile drainage.

• Location and type of conservation practices (buffer strips, filter strips, structural conservation practices such as terracing).

• If operations include livestock, then: (1) livestock inventory; (2) grazing system documentation; (3) manure handling; and (4) location of barns/feeding areas/drainage.

Table 3.1. Some examples of required documentation

BMP	Information/Documentation Required
Nutrient management	The Ohio River program requires three years of farm practice history, including fertilizer application quantities and rate/acre, including fertilizer brand and mixture.
Riparian forest restoration	Current canopy cover, buffer width, stem density, species composition, invasive cover, and channel characteristics (e.g. wetted width). A map with location and extent of BMPs
Cover crop or crop rotation	Last 3 years of crop rotations. A map with location and extent of BMPs
Change in irrigation	Last 3 years of irrigation type, sources of irrigation water (e.g. water diversions, groundwater wells) application rate, and documentation of application. A map with location and extent of BMPs

How confident do we need to be in that documentation?

There is tradeoff between program costs and the level of confidence in documentation of current condition and ability to independently verify those conditions. Comprehensive documentation of conditions will typically improve confidence in baseline values from which credits are calculated and the ability to verify the credit calculation, but at cost to project developers and ultimately, the buyers. Trading programs around the country range widely in the documentation required for current conditions.

Option A *Project developers attest that information provided is true and all available documentation is provided.*	**Option B** *Project developers attest that information is true and provide a required set of documentation.*
Option C *Third parties or agencies collect information and document current conditions.*	

II. Recommended default: Current conditions should be documented using state-approved guidelines provided for each eligible BMP. For structural BMPs, photo points should be included in documenting the current conditions. Project developers collect this documentation and attest that the information is complete and accurate. During verification, the documentation is reviewed for the completeness of information provided.

III. Reasons to deviate from the default: For more complex projects, additional documentation of current condition may be required.

3.2 Initial estimate of project site future conditions

To complete a credit calculation, project developers will also need to project future site conditions after a BMP is installed. For some BMPs, this may be pretty straightforward. For example, current condition is a corn field with no grassed filter strip. Future condition includes a 25-ft wide grassed filter strip and a 1/3 reduction in application of fertilizer. For BMPs that take longer to mature, forecasting future conditions may be a little more challenging. For example, how tall will a riparian forest be and by what date? For either of these scenarios, trading programs need some guidance for project developers on how to estimate and verify post-BMP conditions.

I. Options and examples

In general, there is little formal guidance on how best to forecast future conditions. Oregon DEQ defines future condition of riparian forest buffers as “at maturity.” Willamette Partnership defines future condition as 20 years after BMPs are first implemented, which may not work well for many on-farm BMPs.

<table>
<tr>
<td>Option A
Project developers forecast future conditions based on guidance for each BMP and document their assumptions in a way that can be independently verified.

Pros and Cons?
This approach allows the project developer and their verifier to work out projections of future conditions that are specific to a site. It also provides additional opportunity for conflict over what those assumptions are.</td>
<td>Option B
Trading programs provide explicit guidance on forecasting future conditions (e.g. setting maximum tree heights to be used in model assumptions)

Pros and Cons?
This approach provides a much more repeatable and verifiable way to track projection of future condition. Yet, it is not as sensitive to differences based on site locations.</td>
</tr>
<tr>
<td colspan="2">Option C
Trading programs ask project developers to provide 2-3 future management scenarios that provide a range of possible credit quantities (e.g. under different flow regimes, climate, etc.). Project developers then identify their most likely future scenario for use in generating the credit estimates.

Pros and Cons?
This approach provides a more realistic vision of possible futures. Asking project developers to present scenarios may encourage a thoughtful consideration of whether they should be planning for a change in the quantity of available credits over time. Developing multiple scenarios can also increase the cost of preparing credit estimates.</td>
</tr>
</table>

II. Recommended default: Guidelines for each eligible BMP should include some guidance on what constitutes the future condition. For many BMPs and credit quantification tools, this is simply presence or absence of a particular BMP. For others, project developers will need to document the assumptions built into their estimate of future condition in a way that can be independently verified over time.

III. Reasons to deviate from the default: There may be instances where states choose to review estimates of future condition on a case-by-case basis. Even if trading programs used more fixed assumptions of future conditions, there may be instances where site conditions demand adjusting those fixed assumptions.